

OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/028,384

DATE: 01/16/2002

TIME: 14:42:42

Input Set A:\es.txt

Output Set N:\CRF3\01162002\J028384.raw

ENTERED

```

3 <110> APPLICANT: COMPATIGENE INC.
4   PERREAULT, Claude
5   MCBRIDE, Kevin
7 <120> TITLE OF INVENTION: Mammalian SIMP, gene sequence and uses thereof in cancer
therapy
9 <130> FILE REFERENCE: 5600-74
C--> 11 <140> CURRENT APPLICATION NUMBER: US/10/028,384
C--> 11 <141> CURRENT FILING DATE: 2001-12-20
11 <160> NUMBER OF SEQ ID NOS: 13
13 <170> SOFTWARE: PatentIn version 3.1
15 <210> SEQ ID NO: 1
16 <211> LENGTH: 2481
17 <212> TYPE: DNA
18 <213> ORGANISM: Homo sapiens
20 <220> FEATURE
21 <221> NAME/KEY: CDS
23 <222> LOCATION: (1)..(2481)
24 <223> OTHER INFORMATION
26 <400> SEQUENCE: 1
27 atg gcg gaa ccc tgc gcc ccg gag agc aag cac aag tcg tcc ctc aac      48
28 Met Ala Glu Pro Ser Ala Pro Glu Ser Lys His Lys Ser Ser Leu Asn
29 1 5 10 15
31 tgc tcc ccg tgg agt gcc ctc atg gcc ctg gga aac agc cgg cac gcc      96
32 Ser Ser Pro Trp Ser Gly Leu Met Ala Leu Gly Asn Ser Arg His Gly
33 20 25 30
35 cac cac gga ccc ggg gcc cag tgc gcg cac aag gcg gcg gcc gcc gcg      144
36 His His Gly Pro Gly Ala Gln Cys Ala His Lys Ala Ala Gly Gly Ala
37 35 40 45
39 gcg ccg ccg aag ccc gcc ccg gcg ggg ctg tcc ggg ggg ctg tgc cag      192
40 Ala Pro Pro Lys Pro Ala Pro Ala Gly Leu Ser Gly Gly Leu Ser Gln
41 50 55 60
43 ccg gct gga tgg cag tgc ctt ctc tcc ttc acc atc ctc ttc ctg gcc      240
44 Pro Ala Gly Trp Gln Ser Leu Leu Ser Phe Thr Ile Leu Phe Leu Ala
45 65 70 75 80
47 tgg ctt gcc ggc ttc agc tgc cgc ctc ttc gcc gtc atc cgc ttc gaa      288
48 Trp Leu Ala Gly Phe Ser Ser Arg Leu Phe Ala Val Ile Arg Phe Glu
49 85 90 95
51 agc atc atc cac gag ttc gac ccg tgg ttt aac tat aga tca aca cat      336
52 Ser Ile Ile His Glu Phe Asp Pro Trp Phe Asn Tyr Arg Ser Thr His
53 100 105 110
55 cat ctt gca tct cat ggg ttc tat gaa ttt tta aat tgg ttt gat gaa      384
56 His Leu Ala Ser His Gly Phe Tyr Glu Phe Leu Asn Trp Phe Asp Glu
57 115 120 125
59 aga gca tgg tat cca cta gga aga ata gta ggt ggt act gtt tac cca      432
60 Arg Ala Trp Tyr Pro Leu Gly Arg Ile Val Gly Thr Val Tyr Pro
61 130 135 140
63 ggg ttg atg ata acc gct gcc ctt att cat tgg att tta aat aca ttg      480
64 Gly Leu Met Ile Thr Ala Gly Leu Ile His Trp Ile Leu Asn Thr Leu

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/028,384

DATE: 01/16/2002

TIME: 14 42:42

Input Set : A:\es.txt

Output Set: N:\CRF3\01162002\J028384.raw

65	145	150	155	160	
67	aac ata act gtt cac ata aga gac gta tgt gtg ttc ctt gca cca aat	528			
68	Asn Ile Thr Val His Ile Arg Asp Val Cys Val Phe Leu Ala Pro Thr				
69	165	170	175		
71	ttt ajc ggc ctt aca tct ata tct act ttc ctg ctt aca aga gaa ctt	576			
72	Phe Ser Gly Leu Thr Ser Ile Ser Thr Phe Leu Leu Thr Arg Glu Leu				
73	180	185	190		
75	tgg aac caa gga gca gga ctt tta gct gct tgt ttt att gct att gta	624			
76	Trp Asn Gln Gly Ala Gly Leu Leu Ala Ala Cys Phe Ile Ala Ile Val				
77	195	200	205		
79	cca ggc tac ata tct cgg tca gta gct gga tcc ttt gat aat gaa ggc	672			
80	Pro Gly Tyr Ile Ser Arg Ser Val Ala Gly Ser Phe Asp Asn Glu Gly				
81	210	215	220		
83	att gct att ttt gca ctt cag ttc aca tac tat tta tgg gta aaa tct	720			
84	Ile Ala Ile Phe Ala Leu Gln Phe Thr Tyr Tyr Leu Trp Val Lys Ser				
85	225	230	235	240	
87	gta aaa act ggg tca gtt ttt tgg aca atg tgc tgc tgc tta tcc tat	768			
88	Val Lys Thr Gly Ser Val Phe Trp Thr Met Cys Cys Cys Leu Ser Tyr				
89	245	250	255		
91	ttc tat atg gtc tct gct tgg ggt ggt tat gta ttt atc atc aat ctt	816			
92	Phe Tyr Met Val Ser Ala Trp Gly Gly Tyr Val Phe Ile Ile Asn Leu				
93	260	265	270		
95	att cca ctg cat gta ttt gtg ttg tta ctg atg cag aga tac agc aaa	864			
96	Ile Pro Leu His Val Phe Val Leu Leu Leu Met Gln Arg Tyr Ser Lys				
97	275	280	285		
99	aga gtc tac ata gca tat agc act ttc tac att gtg ggt tta ata tta	912			
100	Arg Val Tyr Ile Ala Tyr Ser Thr Phe Tyr Ile Val Gly Leu Ile Leu				
101	290	295	300		
103	tca atg cag ata cct ttt gtg gga ttc cag cca atc aga aca agt gaa	960			
104	Ser Met Gln Ile Pro Phe Val Gly Phe Gln Pro Ile Arg Thr Ser Glu				
105	305	310	315	320	
107	cac atg gca gct gca ggt gtc ttt gca ttg ctg caa gct tat gct ttc	1008			
108	His Met Ala Ala Ala Gly Val Phe Ala Leu Leu Gln Ala Tyr Ala Phe				
109	325	330	335		
111	ttg cag tat ctg aga gac cga tta aca aaa caa gag ttc cag acc ctt	1056			
112	Leu Gln Tyr Leu Arg Asp Arg Leu Thr Lys Gln Glu Phe Gln Thr Leu				
113	340	345	350		
115	ttc ttt ttg ggt gta tca cta gct gca ggt gct gtg ttc ctt agt gtc	1104			
116	Phe Phe Leu Gly Val Ser Leu Ala Ala Gly Ala Val Phe Leu Ser Val				
117	355	360	365		
119	atc tat ttg act tat aca ggt tac att gca cca tgg agt ggc agg ttt	1152			
120	Ile Tyr Leu Thr Tyr Thr Gly Tyr Ile Ala Pro Trp Ser Gly Arg Phe				
121	370	375	380		
123	tat tca ttg tgg gat act ggg tat gca aaa ata cac att cca att att	1200			
124	Tyr Ser Leu Trp Asp Thr Gly Tyr Ala Lys Ile His Ile Pro Ile Ile				
125	385	390	395	400	
127	gca tca gtg tct gag cat caa cct acg act tgg gtg tct ttc ttc ttt	1248			
128	Ala Ser Val Ser Glu His Gln Pro Thr Trp Val Ser Phe Phe				
129	405	410	415		

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/028,384

DATE: 01/16/2002

TIME: 14:42:42

Input Set : A:\es.txt

Output Set: N:\CRF3\01162002\J028384.raw

131	gat	cta	cat	att	ctt	gta	tgt	acc	ttc	cca	gca	ggc	ctt	tgg	ttc	egg	1296
132	Asp	Leu	His	Ile	Leu	Val	Cys	Thr	Phe	Pro	Ala	Gly	Leu	Trp	Phe	Cys	
133				420					425					430			
135	atc	aaa	aat	atc	aac	gat	gaa	aga	gta	ttt	gtt	gct	cta	tat	gca	atc	1441
136	Ile	Lys	Asn	Ile	Asn	Asp	Glu	Arg	Val	Phe	Val	Ala	Leu	Tyr	Ala	Ile	
137			435					440						445			
139	agt	gct	gtc	tac	ttt	gct	gga	gtg	atg	gtg	cga	ctg	atg	tta	act	ttg	1492
140	Ser	Ala	Val	Tyr	Phe	Ala	Gly	Val	Met	Val	Arg	Leu	Met	Leu	Thr	Leu	
141			450				455							460			
143	act	cca	gtc	gtg	tgt	atg	ctg	tct	gca	att	gcc	ttt	tca	aat	gtt	tct	1449
144	Thr	Pro	Val	Val	Cys	Met	Leu	Ser	Ala	Ile	Ala	Phe	Ser	Asn	Val	Phe	
145	465					470					475				480		
147	gag	cac	tat	ttg	ggt	gat	gac	atg	aaa	agg	gaa	aat	cca	ctt	gtg	gag	1488
148	Glu	His	Tyr	Leu	Gly	Asp	Asp	Met	Lys	Arg	Glu	Asn	Pro	Pro	Val	Glu	
149				485					490					495			
151	gac	acc	aat	gat	gag	gat	gac	aaa	aga	aac	caa	gga	aat	tta	tat	gat	1536
152	Asp	Ser	Ser	Asp	Glu	Asp	Asp	Lys	Arg	Asn	Gln	Gly	Asn	Leu	Tyr	Asp	
153				500				505						510			
155	aag	gca	ggt	aaa	gta	agg	aaa	cat	gca	act	gaa	cag	gaa	aaa	act	gaa	1584
156	Lys	Ala	Gly	Lys	Val	Arg	Lys	His	Ala	Thr	Glu	Gln	Glu	Lys	Thr	Glu	
157			515				520						525				
159	gag	gga	tta	gac	aat	ata	aaa	agc	att	gtc	acc	atg	tta	atg	ctg		1632
160	Glu	Gly	Leu	Gly	Pro	Asn	Ile	Lys	Ser	Ile	Val	Thr	Met	Leu	Met	Leu	
161			530				535						540				
163	atg	cta	tta	atg	atg	ttt	gct	gtc	cac	tgt	acc	tgg	gtc	aca	agc	aat	1680
164	Met	Leu	Leu	Met	Met	Phe	Ala	Val	His	Cys	Thr	Trp	Val	Thr	Ser	Asn	
165	545					550					555				560		
167	gac	tac	tct	agt	cca	agt	gta	gtc	ctg	gcc	tca	tac	aat	cat	gat	ggc	1728
168	Ala	Tyr	Ser	Ser	Pro	Ser	Val	Val	Leu	Ala	Ser	Tyr	Asn	His	Asp	Gly	
169				565					570					575			
171	acc	agg	aat	atc	tta	gat	gat	tta	aga	gaa	gct	tac	ttt	tgg	cta	agg	1776
172	Thr	Arg	Asn	Ile	Leu	Asp	Asp	Phe	Arg	Glu	Ala	Tyr	Phe	Trp	Leu	Arg	
173				580				585						590			
175	caa	aat	aca	gat	gaa	cat	gca	cga	gta	atg	tct	tgg	tgg	gat	tat	ggc	1824
176	Gln	Asn	Thr	Asp	Glu	His	Ala	Arg	Val	Met	Ser	Trp	Trp	Asp	Tyr	Gly	
177			595				600						605				
179	tat	cag	ata	gct	gga	atg	gct	aat	aga	act	acg	tta	gtg	gat	aat	aac	1872
180	Tyr	Gln	Ile	Ala	Gly	Met	Ala	Asn	Arg	Thr	Thr	Leu	Val	Asp	Asn	Asn	
181			610				615						620				
183	acc	tgg	aat	aac	agc	cac	ata	gca	ctg	gtg	gga	aaa	gct	atg	tct	tct	1920
184	Thr	Thr	Trp	Asn	Asn	Ser	His	Ile	Ala	Leu	Val	Gly	Lys	Ala	Met	Ser	
185	625					630					635				640		
187	aat	gaa	aca	gca	gcc	tat	aaa	atc	atg	agg	act	cta	gat	gta	gat	tat	1968
188	Asn	Glu	Thr	Ala	Ala	Tyr	Lys	Ile	Met	Arg	Thr	Leu	Asp	Val	Asp	Tyr	
189				645					650					655			
191	gtt	tta	gtt	att	ttt	gga	ggg	gtt	att	ggc	tat	tct	ggt	gat	gat	atc	2016
192	Val	Leu	Val	Ile	Phe	Gly	Gly	Val	Ile	Gly	Tyr	Ser	Gly	Asp	Asp	Ile	
193				660				665						670			
195	aac	aaa	ttt	ctc	tgg	atg	gtt	agg	ata	gct	gaa	gga	gaa	cat	ccc	aaa	2064

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/028,384

DATE: 01/16/2002

TIME: 14 42:42

Input Set : A:\es.txt

Output Set: N:\CRF3\01162002\J028384.raw

```

196 Asn Lys Phe Leu Trp Met Val Arg Ile Ala Glu Gly Glu His Pro Lys
197          675          680          685
199 gac att cpg gaa agt gac tat ttt acc cca cag gga gaa ttc cgt gta      2112
200 Asp Ile Arg Glu Ser Asp Tyr Phe Thr Pro Gln Gly Glu Phe Arg Val
201      690          695          700
203 gac aaq gaa gga tcc cct act ttg ttg aat tgc ctt atg tat aaa atg      2160
204 Asp Lys Ala Gly Ser Pro Thr Leu Leu Asn Cys Leu Met Tyr Lys Met
205 705          710          715          720
207 tga taa taa aga ttt gga gaa atg cag ctg gat ttt cgt aca ccc cca      2208
208 Ser Tyr Tyr Arg Phe Gly Glu Met Gln Leu Asp Phe Arg Thr Pro Pro
209          725          730          735
211 ggt ttt gac cga aca cgt aat gct gag att gga aat aag gac att aaa      2256
212 Gly Phe Asp Arg Thr Arg Asn Ala Glu Ile Gly Asn Lys Asp Ile Lys
213          740          745          750
215 ttc aaa cat ttg gaa gaa gcc ttt aca tca gaa cac tgg ctt gtt aag      2304
216 Phe Lys His Leu Glu Glu Ala Phe Thr Ser Glu His Trp Leu Val Arg
217          755          760          765
219 ata tat aaa gta aaa gca cct gat aac agg gag aca tta gat cac aaa      2352
220 Ile Tyr Lys Val Lys Ala Pro Asp Asn Arg Glu Thr Leu Asp His Lys
221      770          775          780
223 cct cga gtc acc aac att ttc cca aaa cag aag tat ttg tca aag aag      2400
224 Pro Arg Val Thr Asn Ile Phe Pro Lys Gln Lys Tyr Leu Ser Lys Lys
225 785          790          795          800
227 act acc aaa agg aag cgt gcc tac att aaa aat aag ctg gtt ttt aag      2448
228 Thr Thr Lys Arg Lys Arg Gly Tyr Ile Lys Asn Lys Leu Val Phe Lys
229          805          810          815
231 aaq agc aag aaa ata tct aag aag act gtt taa      2481
232 Lys Gly Lys Lys Ile Ser Lys Lys Thr Val
233          820          825
236 <210> SEQ ID NO: 2
237 <211> LENGTH 826
238 <212> TYPE: PRT
239 <213> ORGANISM: Homo sapiens
241 <240> SEQUENCE: 2
243 Met Ala Glu Pro Ser Ala Pro Glu Ser Lys His Lys Ser Ser Leu Asn
244 1          5          10          15
247 Ser Ser Pro Trp Ser Gly Leu Met Ala Leu Gly Asn Ser Arg His Gly
248      20          25          30
251 His His Gly Pro Gly Ala Gln Cys Ala His Lys Ala Ala Gly Gly Ala
252      35          40          45
255 Ala Pro Pro Lys Pro Ala Pro Ala Gly Leu Ser Gly Gly Leu Ser Gln
256      50          55          60
259 Pro Ala Gly Trp Gln Ser Leu Leu Ser Phe Thr Ile Leu Phe Leu Ala
260 65          70          75          80
263 Trp Leu Ala Gly Phe Ser Ser Arg Leu Phe Ala Val Ile Arg Phe Glu
264          85          90          95
267 Ser Ile Ile His Glu Phe Asp Pro Trp Phe Asn Tyr Arg Ser Thr His
268      100          105          110
271 His Leu Ala Ser His Gly Phe Tyr Glu Phe Leu Asn Trp Phe Asp Glu

```

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/10/028,384

DATE: 01/16/2002
 TIME: 14:42:42

Input Set : A:\es.txt
 Output Set : N:\CRF3\01162002\J028384.raw

```

272          115          120          125
275 Arg Ala Trp Tyr Pro Leu Gly Arg Ile Val Gly Gly Thr Val Tyr Pro
276          130          135          140
279 Gly Leu Met Ile Thr Ala Gly Leu Ile His Trp Ile Leu Asn Thr Leu
280 145          150          155          160
283 Asn Ile Thr Val His Ile Arg Asp Val Cys Val Phe Leu Ala Pro Thr
284          165          170          175
287 Phe Ser Gly Leu Thr Ser Ile Ser Thr Phe Leu Leu Thr Arg Glu Leu
288          180          185          190
291 Trp Asn Gln Gly Ala Gly Leu Leu Ala Ala Cys Phe Ile Ala Ile Val
292          195          200          205
295 Pro Gly Tyr Ile Ser Arg Ser Val Ala Gly Ser Phe Asp Asn Glu Gly
296          210          215          220
299 Ile Ala Ile Phe Ala Leu Gln Phe Thr Tyr Tyr Leu Trp Val Lys Ser
300 225          230          235          240
303 Val Lys Thr Gly Ser Val Phe Trp Thr Met Cys Cys Leu Ser Tyr
304          245          250          255
307 Phe Tyr Met Val Ser Ala Trp Gly Gly Tyr Val Phe Ile Ile Asn Leu
308          260          265          270
311 Ile Pro Leu His Val Phe Val Leu Leu Leu Met Gln Arg Tyr Ser Lys
312          275          280          285
315 Arg Val Tyr Ile Ala Tyr Ser Thr Phe Tyr Ile Val Gly Leu Ile Leu
316          290          295          300
319 Ser Met Gln Ile Pro Phe Val Gly Phe Gln Pro Ile Arg Thr Ser Glu
320 305          310          315          320
323 His Met Ala Ala Ala Gly Val Phe Ala Leu Leu Gln Ala Tyr Ala Phe
324          325          330          335
327 Leu Gln Tyr Leu Arg Asp Arg Leu Thr Lys Gln Glu Phe Gln Thr Leu
328          340          345          350
331 Phe Phe Leu Gly Val Ser Leu Ala Ala Gly Ala Val Phe Leu Ser Val
332          355          360          365
335 Ile Tyr Leu Thr Tyr Thr Gly Tyr Ile Ala Pro Trp Ser Gly Arg Phe
336          370          375          380
339 Tyr Ser Leu Trp Asp Thr Gly Tyr Ala Lys Ile His Ile Pro Ile Ile
340 385          390          395          400
343 Ala Ser Val Ser Glu His Gln Pro Thr Thr Trp Val Ser Phe Phe Phe
344          405          410          415
347 Asp Leu His Ile Leu Val Cys Thr Phe Pro Ala Gly Leu Trp Phe Cys
348          420          425          430
351 Ile Lys Asn Ile Asn Asp Glu Arg Val Phe Val Ala Leu Tyr Ala Ile
352          435          440          445
355 Ser Ala Val Tyr Phe Ala Gly Val Met Val Arg Leu Met Leu Thr Leu
356          450          455          460
359 Thr Pro Val Val Cys Met Leu Ser Ala Ile Ala Phe Ser Asn Val Phe
360 465          470          475          480
363 Glu His Tyr Leu Gly Asp Asp Met Lys Arg Glu Asn Pro Pro Val Glu
364          485          490          495
367 Asp Ser Ser Asp Glu Asp Asp Lys Arg Asn Gln Gly Asn Leu Tyr Asp
368          500          505          510

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/028,384

DATE: 01/16/2002

TIME: 14:42:43

Input Set : A:\es.txt

Output Set: N:\CRF3\01162002\J028384.raw

L:11 M:270 C: Current Application Number differs, Replaced Current Application No

L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date